|  |  |
| --- | --- |
| HSB_Horizontal_RGB **Lösungshinweise**WS 2018/19 | Bildergebnis für strand bar**Prof. Dr. Peter Schmidt****SIB**www.schmidt-bremen.de |

**Statistik EFA**
Montag, 4.2.2019

1. Im **Multiple-Choice-Teil** (Aufg. 1)ergibt jede richtige Lösung (zu einer Aussage) 1,5 Punkte.
Es können hier also insgesamt 21 Punkte (+9 Bonuspunkte=30) erzielt werden.
2. Die Lösungen von **Fill-In-Aufgaben** (Aufg. 2 bis 6) sind in die dafür vorgesehenen Felder (Kästchen) einzutragen. Dies gilt für Skizzen, Grafiken und Texte analog.
**Die Lösungen gelten nur dann, wenn der vollständige Lösungsweg erkennbar ist !**
3. Es ist NUR in der gehefteten Klausur zu arbeiten. Als Konzeptseiten können die Rückseiten der Blätter benutzt werden. **Die Blätter müssen geheftet bleiben**. Bitte benutzen Sie *keinen* Mantelbogen.
4. Es können 90 Punkte\* erreicht werden. Die erreichbare Punktzahl der einzelnen Aufgaben ist angegeben. Da auch die Bearbeitungszeit **90 Minuten** beträgt, sind die Punktzahlen ein Anhalt für die sinnvolle Bearbeitungszeit der Aufgaben. Tipp: Bearbeiten Sie möglichst viele Aufgaben.
*\* + 9+3 Bonuspunkte, die maximale Punktzahl bleibt 90 = 100%.*
5. Zulässige Hilfsmittel:
 - Formelsammlung (P. Schmidt) *ohne* eigene verbale Erläuterungen.
 - Taschenrechner ohne Textverarbeitungsfunktion.
Bitte schreiben Sie nicht mit roten Stift und benutzen Sie Bleistifte nur in Zeichnungen.

***Viel Erfolg !***

1. **Mobiltelefone** (sowie **Smartwatches** u.a.) **sind verboten.** Schon Bereithalten **am Platz** gilt als Täuschungsversuch  in eine Tasche und diese entfernt vom Sitzplatz deponieren**!**
2. Überprüfen Sie zu Beginn die Klausur auf Vollständigkeit (6 Aufgaben auf 9 Seiten) und füllen
 bitte vorab die unten stehenden Kästchen aus *(„Versuch“ = ich schreibe die Klausur zum ... Mal)*. ⮷

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Name**: |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 1. Versuch:  |  |
| **Vorname**: |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 2. Versuch: |  |
| **MatrikelNr:** |  |  |  |  |  |  |  |  |  | Wenn zutreffend, **unbedingt** ankreuzen → **3.** Versuch: |  |

*Aufg. Punkte:*  ***Ab hier bitte nichts beschriften oder ankreuzen:***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1.  | *(21+9)* |  |  |  |
| 2. | *(20)* |  | Datum:  |
| 3. | *(10)* |  |
| 4. | *(20)* |  | Prozent /  |  |
| 5. | *(10 +3)* |  | Note |  |
| 6. | *(9)* |  | Unterschrift |
| **** | ***(max. 90)*** |  |

**Achtung: Bitte keinen Leistungsnachweis-Bogen beilegen!**

Bildquelle: www.strandbar-freiburg.de

 Die **Strandbar**
In Vorfreude auf das Sommersemester gründet der EFA-Studiengang einer norddeutschen Hochschule am Weserstrande eine Bar. In allen Pausen und nach Veranstaltungsschluss bieten sie Eis, Snacks, Cocktails und andere Getränke an.

# **Multiple Choice** Bitte kreuzen Sie an - Erläuterungen sind nicht erforderlich. Hinweis: Eine Multiple-Choice-Aussage ist nur „richtig“, wenn die Aussage *immer* gilt. Gibt es ein einziges Gegenbeispiel, so ist sie „falsch“. *[je 1,5 Punkte ⇒ Gesamt 21+9 Punkte]*

|  |  |  |
| --- | --- | --- |
| richtig | falsch |  Welche Aussage ist richtig? |
|  | x | 1. Mit einer Statistik des Bierkonsums kann man den Zusammenhang zwischen Klausurfrust und Trinkfestigkeit und der Studierenden statistisch beweisen.
 |
| x |  | 1. Ein Rang-Korrelationskoeffizient für Snacks kann negative Werte annehmen.
 |
| x |  | 1. W(Cocktail ∪ Bier) = W(Cocktail) + W(Bier) gilt für sich ausschließende Ereignisse
 |
|  | x | 1. Die Anzahl der möglichen Anordnungen (Reihenfolgen) von Chipstüten auf Theke bezeichnet man als Kombinationen.
 |
| x |  | 1. Aus den Werten einer Ratingskala für Cocktails kann ein Median berechnet werden
 |
|  | x | 1. Es gibt drei Typen von Zufallsvariablen
 |
| x |  | 1. Die Anzahl der verkauften Snacks pro Tag ist eine Zufallsvariable
 |
|  | x | 1. Der Zentrale Grenzwertsatz beschreibt den Mittelwert der Standardfehler
 |
|  | x | 1. Die Ermittlung von Durchschnittsnoten als arithmetisches Mittel ist statistisch korrekt.
 |
| x |  | 1. Ein Zweistichproben-Test untersucht z.B., ob zwei Kundengruppen Strandbar sich statistisch signifikant unterscheiden
 |
| x |  | 1. Die Anzahl der Cocktails, die jemand pro Tag trinkt ist diskret und stetig.
 |
| x |  | 1. Der Korrelationskoeffizient hat dasselbe Vorzeichen wie die Steigung b.
 |
|  | x | 1. Lagemaße sind z.B. Mittelwerte, DAA, Quantile …
 |
| x |  | 1. Ein Konfidenzintervall ist umso schmaler, je größer  ist.
 |
|  | x | 1. Beim Umbasieren und Verketten von Indices wird deren Berichtsperiode geändert
 |
|  | x | 1. Ein Preisindex ist ein gewichtetes geometrisches Mittel aus Preissteigerungsziffern.
 |
| x |  | 1. Durchschnittsnoten werden in Deutschland in der Regel als arithmetisches Mittel gebildet.
 |
|  | x | 1. Ein Cocktail-Preisindex von 120 für 2019 zur Basis 2015 bedeutet, dass die Preise im Durchschnitt um 5 % pro Jahr gestiegen sind.
 |
| x |  | 1. Wenn alle Residuen gleich 0 sind, wird das Bestimmtheitsmaß gleich 1.
 |  |
| x |  | 1. Ob ein Cocktail teurer ist, wenn er mehr Alkohol enthält, kann mit einer Korrelationsanalyse untersucht werden.
 | *(21+9)* |

# Sie analysieren Sie die Umsatzzahlen der Strandbar aus den letzten 3 Jahren. Es wurden jeweils Dritteljahre (Tertiale) aufgezeichnet. Sie sehen, dass die Entwicklung einer saisonalen Entwicklung folgt und möchten eine **saisonbereinigte Umsatzentwicklung** betrachten können. Für die Glättung verwenden Sie eine lineare Trendfunktion. Zusätzlich möchten Sie die zu erwartende Entwicklung für 2019 **prognostizieren.** *(Die Umsatzzahlen sind in TEuro angegeben, aber das brauchen Sie bei der Rechnung nicht zu berücksichtigen. Es ist ausreichend, wenn Sie mit einer Nachkommastelle rechnen)*  *[Gesamt: 20 Punkte]*



Hinweise: Bei Verwendung von einfachen t-Werten (t=1,2,3, ...) ergibt sich **a = 20,3**

## Analysieren Sie die Entwicklung der Umsatzzahlen in den drei Jahren mit einem linearen Trend. Was bedeutet der gefundene Wert von b? Bewerten Sie diese Entwicklung. [5 Punkte]

s.o.: y^ = 30 + 1,93 \* ti\*  [bzw. y^ = 20,3 + 1,93 \* ti]

Positiver Trend, mehr Umsatz, gute Entwicklung

## Ermitteln Sie auf dieser Basis die durchschnittlichen Saisonkomponenten für die Tertiale. Interpretieren Sie die „Saisonfigur“; in welchem Tertial wird am meisten verkauft? Haben Sie eine Idee, warum? [5 Punkte]

s.o. 2. Tertial – Sommer …, Sonne, Strand, Zeit und überstandene Klausuren feiern

## Ermitteln Sie die saisonbereinigte Zeitreihe. Welchen Sinn hat eine solche Saisonbereinigung? [3 Punkte]

s.o. Wie der Name schon sagt, das Herausrechnen der immer wiederkehrenden saisonalen Einflüsse, damit die wirkliche (reale) Entwicklung des Marktes sichtbar wird.

## Zeichnen Sie die Beobachtungswerte, die Gerade und die Werte aus Aufgabe 2.3 [4 Pkt]

 *(Zeichnung s.o.)*

## Ermitteln Sie die zu erwartenden Eintrittszahlen für die drei Tertiale 2019. Berücksichtigen Sie dabei die Saisoneinflüsse. Erläutern Sie Ihr Vorgehen in Stichworten. [3 Punkte]

s.o.

Es ist zu beachten, dass die oben ermittelten durchschnittlichen SK *addiert* werden, da Sie den reinen Trend überlagern. Nur so kann eine realistische Vorhersage der Entwicklung getroffen werden.

*(20)*

# Die durchschnittliche KundInnenzahl sei als Zufallsvariable X  *N* (70,2) verteilt. *[Gesamt: 10 Punkte]*

## Zeichnen (skizzieren) Sie die Dichtefunktion mit einer x- und einer z-Achse (alle Achsen beschriften!). Zeichnen Sie die Werte 67, 68, 73 und 74 auf der x-Achse und die entsprechenden Werte auf der z-Achse ein. [4 Punkte]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| x | 67 | 68 | 73 | 74 |
| z | -1,5 | -1 | 1,5 | 2 |

## Zeichnen Sie die zugehörige Verteilungsfunktion. Markieren Sie den Funktionswert für 74 KundInnen auf der senkrechten Achse. [3 Punkte]



Es ist lediglich eine Skizze gefragt, keine exakten f(x) Werte

Dieser Bereich konnte für Aufgabe 3.3 markiert werden.

Fläche unter der f(x) Kurve: F(x=74)
= F(z = 2)
= **0,9772**)

## Was bedeutet der unter Aufgabe 3.2 gefundene Wert (inhaltliche Aussage)? [3 Punkte]

Die Wahrscheinlichkeit, dass die Zufallsvariable X höchstens des Wert x=74 annimmt, d.h. die Wahrscheinlichkeit, dass höchstens 74 KundInnen anwesend sind.

*(10)*

# **Aber bitte mit Schuss!** Ein wesentliches Qualitätskriterium für den Cocktail **Vögel am Strand** ist der Alkoholgehalt. Um eine hochwertige Produktqualität sicherzustellen, achtet die Strandbar verstärkt auf dieses Kriterium.  *[Ges: 20 Pkt]*

## Für einen Vögel-Cocktail ist es wichtig, dass der Alkoholgehalt von 15 % weder unterschritten (Geschmack) noch überschritten wird (Führerschein). Sie untersuchen 625 Proben (Prost!) und ermittelt einen Alkoholgehalt von 20 %. Kann die Strandbar das Produkt so verkaufen, wenn eine Irrtumswahrscheinlichkeit von 5 % gilt? [7 Punkte]

|  |  |  |  |
| --- | --- | --- | --- |
| Informationen: | Beobachtungen | Mittelwert | Verteilung |
| Grundgesamtheit | k.A. | **po = 0,15** (Aufgabe) | Normalverteilung wird angenommen |
| Stichprobe | 625 | **= 0,20**  |

1) Hypothesen: Ho: p = p0 = 0,15 ; H1: p ≠ p0 → zweiseitiger Test

2)  = 0,05 lt. Aufgabenstellung

3) Keine Endlichkeitskorrektur (da kein N angegeben, als „sehr groß“ angenommen) →
 $σ\_{p\^}\sqrt{\frac{0,2\*0,8}{625}}$ = **0,016**

4) zc = 1,96

5) zp = 3,13 [Alternative B] **pcu = 0,119 pco = 0,181**

6) |zp| > |zc| ? JA! → H0 VERWERFEN ** > pcu aber  > pco → H0 VERWERFEN**

7) Mit einer Wahrscheinlichkeit von 95% ist der Alkoholgehalt NICHT 15% (sondern größer) und daher sollten die Mixer angehalten werden, weniger Alkohol einzurühren.

## Schätzen Sie auf Basis o.a. Stichprobe, wie groß der Alkoholgehalt der Cocktails in der Grundgesamtheit ist.  Was heißen diese Werte inhaltlich? (=auf Deutsch in Ihren Worten) Finden Sie die Schätzung plausibel? Wann wäre der Bereich kleiner? [3 Punkte]

 *Konfidenzintervall*  *W (0,20 –* 1,96*⋅* 0, 016*≤ p ≤ 0,20 +* 1,96*⋅* 0, 016*) = 0,95
 W (****0,169*** *≤  ≤* ***0,231****) = 0,95%*

*Mit 95-prozentiger Wahrscheinlichkeit liegt der Alkoholgehalt zwischen 16,9 und 23,1 Prozent. Dies ist eine Schätzung des unbekannten Anteilswertes der Grundgesamtheit auf Basis eines (zufälligen) Stichprobenanteils.*

*Das ist ein ziemlich breites Intervall und für die Entscheidung nicht besonders hilfreich.
Bei höherem  würde das Intervall kleiner, ebenso bei einer größeren Stichprobe.*

## Skizzieren Sie dieses Intervall zeichnerisch*(natürlich müssen Sie die Achsen und die Werte beschriften!)* [3 Punkte]

 *f(**)*

 *gu  go*

 *0,169*  *0,231* 

## Als Statistiker möchten Sie gerne den Altersdurchschnitt der KundInnen wissen. Vor einigen Jahren betrug dieser 25 Jahre bei einer Standardabweichung von 4; die Population sei normalverteilt. Es wird eine Stichprobe von 64 KundInnen untersucht, deren Durchschnittsalter beträgt 24,2 Jahre. Hat sich das Durchschnittsalter verändert, wenn ein Konfidenzniveau von 95% angenommen wird? [7 Punkte]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Informationen | Beobachtungen | Mittelwert | Standardabweichung | Verteilung |
| Grundgesamtheit | k.A. | Ausgangsannahme: 0=25 | 4 | normalverteilt |
| Stichprobe | 64 | = 24,2 | - |  |

1) Hypothesen**: Ho: μ = μo = 25** und H1: μ ≠ μo → **zwei**seitig kritischer Bereich

2)  = 1 % lt. Aufgabenstellung

3)  bekannt, Normalverteilung **→ 1. Fall** **→ = 4/8 = 0,5**

4) F(zc) = 1-  ⇒ **zc = 1,96**

**5) zx = -1,6 oder cu = 24,020 ; co = 25,980**

**6) |zx| > |zc| → H0 NICHT verwerfen**  **> co → H0 NICHT verwerfen**

7) Das Durchschnittsalter hat sich nicht statistisch signifikant verändert.

*(20)*

# Die Strandbar befragt zwei KundInnen (Peter und Paula), welchen Cocktail sie am liebsten mögen. Peter vergibt Punkte, Paula nennt ihre Lieblings-Reihenfolge. *[ges: 10 P]*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Peters |  | Paula |  |
|  | Punkte |  | Rangfolge |  |
| 7 up the beach | 20 |  | 4 |  |
| Vögel am Strand | 15 |  | 5 |  |
| Mochti-so | 5 |  | 3 |  |
| Kur-a-Sau | 18 |  | 2 |  |
| 7 down to beach | 7 |  | 1 |  |
|  |  |  |  |  |

## Gibt es einen Zusammenhang zwischen den Bewertungen unserer Longdrink-Fans? Welches ist das geeignete Maß, die Stärke des Zusammenhanges zu ermitteln? [3 Punkte]

|  |
| --- |
| Geeignet ist der Rangkorrelationskoeffizient nach SpearmanGrund: ordinal skaliert |

## Errechnen Sie dieses Maß und interpretieren Sie das Ergebnis. [4 Punkte]

|  |
| --- |
| Schwacher negativer Zusammenhang = die Geschmäcker der beiden sind etwas unterschiedlich |

## Welche anderen Zusammenhangmaße gibt es? Beschreiben Sie diese [3 Punkte]

|  |
| --- |
| Hier sollen die verschiedenen Maße in Abhängigkeit von den jeweiligen Skalen beschrieben werden |

BONUS:

## Sie schreiben eine Hausarbeit über die Strandbar und sollen den zeitlichen Verlauf des Absatzes von Bock’s Bier statistisch beschreiben. Welche statistischen Methoden könnten Sie (über die oben schon durchgeführte Regressionsanalyse hinaus) verwenden. [3 Punkte]

|  |
| --- |
| … Bonus …  |

# Gleitende Durchschnitte [Gesamt: 9 Punkte]

## Sie beobachten die Anzahl der KundInnen um 20.30 Uhr an sechs aufeinander folgenden Tagen. Da die Beobachtungswerte stark schwanken, möchten Sie diese glätten. Verwenden Sie dazu Gleitende Durchschnitte dritter und vierter Ordnung[[1]](#footnote-1). [5 P]

 *(Zeichnung s.u.)*

## Zeichnen Sie die beiden gleitenden Durchschnitte in die u.a. Grafik ein und erläutern Sie kurz den „Informationsverlust am Rand“. [4 P]



*(9)*

1. Es ist freigestellt, ob Sie zentrierte oder endwertige GD verwenden. [↑](#footnote-ref-1)